Capacities of Quantum Error Correcting Codes under adaptive concatenation

Jesse Fern
Berkeley Center for Quantum Information and Computation,
Department of Mathematics, University of California, Berkeley, California, 94720 *

We look at the effects of a quantum channel after each level of quantum error correcting codes (QECC) under recovery operators that are optimally adapted at each level. We use the entropy of the channel to estimate the capacities of QECCs. Considerable improvements in capacities are found under adaptive concatenation.

This paper extends the quantum error correcting channel map formalism [6] [4]. We examine how recovery operators can be made optimal. We then find the capacity of different codes under different types of noise, and compare it to the Shannon limit. To calculate the capacity, we use the entropy of the channels instead of previous non-entropy based methods [5].

Basic notation The Pauli operators on n qubits are $\mathcal{P}_n = \{I, X, Y, Z\}^{\otimes n}$. We use the notation that $XY = X \otimes Y$. Each pair of Pauli operators σ, σ' either commutes or anti-commutes. We let $\eta(\sigma, \sigma')$ be 1 if they commute, -1 if they anti-commute.

We write a density matrix on n qubits as $\rho = \frac{1}{2^n} \sum_{\sigma \in \mathcal{P}_n} \rho_\sigma \sigma$. A quasi-density matrix ρ is an unnormalized density matrix. It can be normalized to a density matrix $\frac{1}{p}\rho$, where $p = \text{tr } \rho = \rho_I$.

A channel is a map from density matrices to density matrices, and can be written as $\$(\rho) = \sum_i A_i \rho A_i^\dagger$, where the A_i are Kraus operators that act on the state. This channel has a matrix representation $\mathcal{N} = \sum_i \int(A_i)$, where $\int(M) = M \otimes \overline{M}$, and $\mathcal{N}$ is a matrix that acts on the vector form of ρ in the standard basis.

By looking at how the 4^n Pauli operators $\sigma \in \mathcal{P}_n$ perform under the map $\$(\sigma)$, the channel can be written as a 4^n by 4^n superoperator in the Pauli basis. Because channels are trace-preserving and hermitian-preserving, this implies that 1-qubit channels written in the Pauli basis have real entries and the first row is $\mathcal{N}_{I\sigma} = \delta_{I,\sigma}$. On 1-qubit, this is

$$\mathcal{N}^{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ N_{XI} & N_{XX} & N_{XY} & N_{XZ} \\ N_{YI} & N_{YX} & N_{YY} & N_{YZ} \\ N_{ZI} & N_{ZX} & N_{ZY} & N_{ZZ} \end{bmatrix}.$$

Just as a density matrix can be written as a sum of quasi-density matrices, a channel can be written as a sum of quasi-channels. Quasi-channels map density matrices to quasi-density matrices, and therefore aren't trace preserving. If $\mathcal{N}$ is a quasi-channel, and ρ is a density matrix, then the probability of the quasi-channel is $p = \text{tr}(\mathcal{N}\rho) = \sum_\sigma N_{I\sigma}\rho_\sigma$, which depends on ρ unless $\mathcal{N}_{I\sigma} = p\delta_{I,\sigma}$, in which case this gives the channel $\frac{1}{p}\mathcal{N} = \frac{1}{N_I I}\mathcal{N}$ with probability p.

Overview of Stabilizer codes An $[[n, k, d]]$ quantum stabilizer code encodes k logical qubits into n physical qubits, and has a distance d. Typically, a quantum stabilizer code is given in terms of $n-k$ generators g_i, and 4^k encoded Pauli operators $\overline{\sigma}$.

The g_i generate the stabilizer group S, which is isomorphic to Z_2^{n-k}. The k logical qubits encode into the 2^k dimensional codespace C_S, which has the property that if $|\psi\rangle \in C_S$ and $s \in S$, then $s|\psi\rangle = |\psi\rangle$. The elements of the set $\mathcal{C}(S) \subset \mathcal{P}_n$ are the 2^{n+k} Pauli operators that commute with S. They send states in C_S to states in C_S. Each of the 4^k different equivalence classes of $\mathcal{C}(S)/S$ correspond to one of the 4^k encoded logical Pauli operators $\overline{\sigma}$, which act on the logical qubits in the codespace.

To perform error detection we measure each of the $n-k$ generators g_i, projecting into either the $+1$ or -1 eigenspace for each one, which gives us one syndrome bit β_i, which is 0 if in the $+1$ eigenspace, and 1 if in the -1 eigenspace. These β_i form the 2^{n-k} possible error syndromes β. $\beta = 0 = (0, 0, \ldots, 0)$ is the syndrome corresponding to no error.

To recover from an error with syndrome β, a recovery operator $R(\beta)$ is chosen which returns to the codespace. It has the property that $\eta(R(\beta), g_i) = \beta_i$, and the 2^{n+k} possible choices for $R(\beta)$ are all in the same equivalence class of $\mathcal{P}_n/\mathcal{C}(S)$.

Channel map formalism We define a 4^n by 4^k encoding operator $\mathcal{E}$ that maps from a density matrix on k qubits to one on n qubits as

$$\rho \to \frac{1}{2^{n-k}}\mathcal{E}\rho.$$

This projects us into the codespace of the code. In the Pauli basis, the 4^k columns of $\frac{1}{2^{n-k}}\mathcal{E}$ act as the logical Pauli operators in the code space, and 0 outside of the code space. They are defined as

$$\mathcal{E}_I = \prod_i (I + g_i) \qquad\qquad \mathcal{E}_\sigma = \overline{\sigma}\mathcal{E}_I. \tag{1}$$

Now that the logical qubits are encoded into the codespace, assume that a noise $\mathcal{N}$ acts on the full 2^n dimensional space of the n physical qubits of the code. The g_i operators are measured, giving an error syndrome β. Some recovery operator $R(\beta)$ is then chosen. We can write the error correction process as

$$\mathcal{T}_{\text{ideal}} = \sum_\beta \int(R(\beta)) \circ \mathcal{P}_\beta = \mathcal{P}_0 \circ \sum_\beta \int(R(\beta))$$

where $\mathcal{P}_\beta$ is a projection superoperator into the syndrome β space, and $\mathcal{P}_0 = \frac{1}{2^{n-k}}\mathcal{E} \circ \mathcal{E}^t$ is a projection into the codespace.

$\mathcal{E}^t$ decodes back back into the original k qubit space. Then, the whole process can be written as

$$\mathcal{G} = \frac{1}{2^{n-k}}\mathcal{E}^t \circ \sum_\beta \int(R(\beta)) \circ \mathcal{N} \circ \mathcal{E} \tag{2}$$

This represents noise on the logical qubits $\mathcal{G}$ in terms of the noise on the physical qubits $\mathcal{N}$. We can decompose it as a sum

$$\mathcal{G} = \sum_\beta \mathcal{G}^{R(\beta)}, \mathcal{G}^\sigma = \frac{1}{2^{n-k}}\mathcal{E}^t \circ \int(\sigma) \circ \mathcal{N} \circ \mathcal{E}. \tag{3}$$

of the contributions from each error syndrome.

Same noise on each qubit Often it is assumed that there is the same noise on each qubit, so $\mathcal{N} = \left(\mathcal{N}^{(1)}\right)^{\otimes n}$.

In previous work [6] [4], it was shown that in the case where $k = 1$ logical qubit, this will produce a map
$\Omega^C : \mathcal{N}^{(1)} \to \sum_\beta \frac{1}{2^{n-k}}\mathcal{E}^t \circ \int(R(\beta)) \circ \left(\mathcal{N}^{(1)}\right)^{\otimes n} \circ \mathcal{E}$
from the physical 1-qubit noise on each qubit to the logical 1-qubit noise on the encoded codespace. This is useful for analyzing a code concatenated with itself an arbitrary number of times. By encoding each logical qubit on one level of the code into a physical qubit at the next level of the code, we can concatenate a code with itself many times. In particular, the code will correct noise if $\lim_{k\to\infty} \Omega^{C^{\circ k}}(\mathcal{N}^{(1)}) = I$. However, this assumes that the recovery operators don't depend on syndrome information from the previous levels of error correction, and so don't utilize the full capacity of the code that adaptive concatenation uses.

Diagonal noise A Pauli matrix σ has the superoperator $\int(\sigma)$, which is diagonal, and $\int(\sigma)_{\sigma'\sigma'} = \eta(\sigma, \sigma')$.

Suppose the noise is diagonal in the Pauli basis. Then the $\mathcal{G}^{R(\beta)}$ from Eq. 3 become

$$\mathcal{G}^{\sigma'}_{\sigma,\sigma} = \frac{1}{2^{n-k}} \sum_{s\in S} \eta(\sigma', \overline{\sigma}s)\mathcal{N}_{\overline{\sigma}s,\overline{\sigma}s}. \tag{4}$$

If a quasi-channel is diagonal, it can be written as a sum of Pauli superoperators $\mathcal{N} = \sum_\sigma p_\sigma \int(\sigma)$, which corresponds to having the Pauli error σ with probability p_σ. For a 1-qubit quasi-channel with probability $p = \sum_\sigma p_\sigma$, the diagonal parts are

$$\begin{aligned}[p, x, y, z] = p_I[1, 1, 1, 1] + p_X[1, 1, -1, -1]\\ +p_Y[1, -1, 1, -1], +p_Z[1, -1, -1, 1].\end{aligned} \tag{5}$$

Then the Pauli probabilities are

$$p_I = \frac{p + x + y + z}{4} \qquad p_X = \frac{p + x - y - z}{4}$$
$$p_Y = \frac{p - x + y - z}{4} \qquad p_Z = \frac{p - x - y + z}{4}.$$

Because the Shannon entropy in terms of Pauli errors q_i is $\sum_i f_s(q_i)$ where $f_s(x) = -x \log_2 x$, a quasi-channel gives a contribution of $p \sum_\sigma f_s(\frac{p_\sigma}{p}) = -f_s(p) + \sum_\sigma f_s(p_\sigma)$ to the entropy of a channel.

Adaptive concatenation In this section, we discuss using syndrome information from a code on the lower levels of a code, and how this can be used to improve the threshold of a code. The results here are more general than previous work [5].

The logical encoded channel from a code can written as a sum $\mathcal{G} = \sum_\beta G^{R(\beta)} \otimes \beta$ over the contribution to the channel from each error syndrome $\mathcal{G}^{R(\beta)}$, which is given in Eq. 3.

If we perform error correction, and ignore the syndrome information, then this sum just becomes $\mathcal{G}$, but if we know we had the syndrome β, we have the quasi-channel $\mathcal{G}^{R(\beta)}$. The probability of this syndrome occurring depends on the density matrix state ρ and is

$$p_\beta = tr(\mathcal{G}^{R(\beta)}\rho) = 2^n \sum_\sigma \mathcal{G}^{R(\beta)}_{I\sigma}\rho_\sigma.$$

If the first row of this quasi-channel in the Pauli basis is all zero except for the first term, that is $\mathcal{G}^{R(\beta)}_{I\sigma} = p\delta_{I\sigma}$, then $p_\beta = \mathcal{G}^{R(\beta)}_{II} = p$, and measuring the syndrome β causes us to collapse the channel $\frac{1}{p_\beta}\mathcal{G}^{R(\beta)}$.

Recovery optimization Given an $[[n, 1, d]]$ quantum code, assume that the syndrome β was measured. We choose some representative element r_β of that equivalence class of $\mathcal{P}_n/C(S)$ to return to the codespace. This differs from the optimal recovery operator $R(\beta)$ by a logical Pauli operator $\overline{\sigma}$. The recovery operators for that syndrome are then equivalent to the 4 recovery operators $\overline{\sigma}r_\beta$, where $\overline{\sigma}$ is any one of the 4 logical encoded Pauli operators.

Now that we have a quasi-channel corresponding to the error syndrome that was measured, we pass this information onto the next level of concatenation. For each of the n blocks of the next level, we will have a (most likely) different quasi-channel $\mathcal{N}^{(1)}_i$. The combined noise is then the quasi-channel $\mathcal{N} = \bigotimes_i \mathcal{N}^{(1)}_i$. We must optimize the recovery function based upon these quasi-channels.

There are two methods of optimization. The first method involves finding the optimal recovery operator $r_\beta\overline{\sigma}$ at each level of the code based upon the noise input.

The second method involves just applying some random recovery operator r_β at each level of the code, instead of the optimal recovery operator. Since $\mathcal{G}^{\sigma r_\beta} = \int(\sigma)\mathcal{G}^{r_\beta}$, this causes a difference of a logical Pauli operator at that level of the code, and a difference of a Pauli

TABLE I: Encoded Pauli operators $\overline{\sigma}$ and encoding operator $\mathcal{E}_\sigma$ for the 2 qubit code

σ	$\overline{\sigma}$	$\mathcal{E}_\sigma$
I	II	$II+ZZ$
X	XX	$XX-YY$
Y	XY	$XY+YX$
Z	IZ	$IZ+ZI$

operator $\sigma' \in \mathcal{P}_n$ at the next level of the code. At the next level of the code, the noise is $\mathcal{N}' = \int(\sigma')\circ\mathcal{N}$ instead of $\mathcal{N}$. If we apply the recovery operator r_β, we get $\mathcal{G}^{r_\beta\sigma'} = \frac{1}{2^{n-k}}\mathcal{E}^t\circ\int(r_\beta)\circ\int(\sigma')\circ\mathcal{N}\circ\mathcal{E}$ instead of $\mathcal{G}^{r_\beta}$. Now, $r_\beta\sigma' = \overline{\sigma''}r_\alpha$ for some syndrome α and a logical Pauli operator $\overline{\sigma''}$, so $\mathcal{G}^{r_\beta\sigma'} = \int(\overline{\sigma''})\mathcal{G}^{r_\alpha}$, and so this causes a difference of a logical Pauli operator at the next level of the code.

While the first method may seem more intuitive, the second may be easier computationally. It might not be clear what the optimal recovery operator at some lower level of the code is. At the last level of the code, we choose the encoded Pauli operator that gives the channel that is closest to the identity channel. If the resulting noise is very close to the identity channel, then very little logical noise is introduced at each step of a quantum computation, allowing long quantum computations to be performed. Oftentimes, the noise at the end will be close to diagonal even though the initial noise was not.

Suppose we wish to calculate the exact optimized channel map of a $[[n,1,d]]$ quantum code after j levels of concatenation with itself. There are 2^{n^j-1} syndromes. Many of these syndromes may have the same channel components $\mathcal{G}^{R(\beta)}$, so the complexity involved might be less than this, but it is still too computationally expensive to compute the optimized channel for more than the 2nd level of many codes.

However, the overhead involved with actually implementing the optimization is low, because the encoded channel from one level is just passed on to the qubits of the next level of the code. This leads to efficient Monte Carlo simulation.

Example: The 2 qubit bit flip code The 2 qubit bit flip code is useful to look at because it is so simple compared to other codes.

The code is the classical 2 bit majority code. Its codespace is spanned by the logical encoded states $\overline{|0\rangle} = |00\rangle$, and $\overline{|1\rangle} = |11\rangle$. The Stabilizer group $S=\{II, ZZ\}$ is generated by the one generator $g_1 = ZZ$. In Tab. I, we have the encoded Pauli operators, and from Eq. 1 calculate the $\mathcal{E}_\sigma$.

Using Eq. 3, we can calculate various channel map components $G^{R(\beta)}$. For example, $\mathcal{G}^{IX}_{X,Z} = \frac{1}{2}(-N_{XX,IZ} + N_{XX,ZI} + N_{YY,IZ} - N_{YY,ZI})$. In the case of diagonal noise, the channel map components can be found directly

TABLE II: $\mathcal{G}^{R(\beta)}$ for diagonal noise, using the shorthand $\sigma = N_{\sigma,\sigma}$

β	$R(\beta)$	$\mathcal{G}^{R(\beta)}$
0	II	$\frac{1}{2}[II+ZZ, XX+YY, XY+YX, IZ+ZI]$
1	XI	$\frac{1}{2}[II-ZZ, XX-YY, XY-YX, IZ-ZI]$
1	IX	$\frac{1}{2}[II-ZZ, XX-YY, YX-XY, ZI-IZ]$

from Eq. 4, and are given in Tab. II, using the shorthand $\mathcal{N}_{\sigma,\sigma} = \sigma$.

Suppose now that we have the same diagonal noise $\mathcal{N}^{(1)} = [1,1,x,x]$ on each qubit, with $x>0$. We see from Eq. 5 that this represents an X error with probability $p_x = \frac{1-x}{2}$. Then,

$$\mathcal{G}^{II} = [\frac{1+x^2}{2}, \frac{1+x^2}{2}, x, x]$$

$$\mathcal{G}^{IX} = \mathcal{G}^{XI} = [\frac{1-x^2}{2}, \frac{1-x^2}{2}, 0, 0]$$

The total map is given by

$$\Omega^{bf_2}[1,1,x,x] = \mathcal{G} = \mathcal{G}^{II} + \mathcal{G}^{IX} = \mathcal{G}^{II} + \mathcal{G}^{XI} = [1,1,x,x],$$

and so this code isn't very useful at correcting bit flip errors.

Suppose that we concatenate the code with itself. It acts on the bit flip channel as $\Omega^{bf_2}\circ\Omega^{bf_2}[1,1,x,x] = [1,1,x,x]$, which is again useless. The problem is that the recovery operators are also concatenated using this method. Let the recovery operators be $R=\{II, XI\}$. Now, we wish to optimize the recovery operator on the second level for the noise. The recovery operator was already optimal unless there was no error detected in the first block, which gives a block qubit noise of $\mathcal{N}_1 = \mathcal{G}^{II} = [\frac{1+x^2}{2}, \frac{1+x^2}{2}, x, x]$, and there was an error detected in the second block, which gives a block qubit noise of $\mathcal{N}_2 = \mathcal{G}^{XI} = [1,1,x,x]$, and there was an error detected at the top level of the code, which gives a logical noise of

$$\begin{aligned}
\mathcal{G}^{XI} &= \frac{1}{2}[II-ZZ, XX-YY, XY-YX, IZ-ZI] \\
&= \frac{1}{2}[\frac{1+x^2}{2}\frac{1-x^2}{2} - x0, \frac{1+x^2}{2}\frac{1-x^2}{2} - x0, \\
&\quad \frac{1+x^2}{2}0 - x\frac{1-x^2}{2}, \frac{1+x^2}{2}0 - x\frac{1-x^2}{2}] \\
&= \frac{1-x^4}{8}[1,1,-\frac{2x}{1+x^2}, -\frac{2x}{1+x^2}].
\end{aligned}$$

This is the same as the channel component $\mathcal{G}^{XXXI}$ for the 4 qubit bit flip code.

Since the last 2 components are always negative, it would be optimal to apply the recovery operator IX instead, since it differs by encoded $\overline{X} = XX$. It gives

$$\mathcal{G}^{IX} = [1,1,-1,-1]\mathcal{G}^{XI} = \frac{1-x^4}{8}[1,1,\frac{2x}{1+x^2}, \frac{2x}{1+x^2}],$$

TABLE III: Critical values for depolarizing channel $(p_X, p_Y, p_Z) = (p, p, p)$

Level	[[5,1,3]]	[[7,1,3]]	Special code
0	6.30965616%	6.30965616%	6.30965616%
1	6.29873094%	6.25921455%	6.34520294%
2	6.29795843%	6.26714580%	6.35204743%
3	6.298(5)%	6.268(8)%	6.368(1)%
4	6.299(0)%	6.269(6)%	6.367(2)%
5	6.299(3)%	6.270(0)%	6.367(8)%
6	6.299(5)%	6.270(3)%	6.368(3)%
7	6.299(6)%	6.270(3)%	6.368(5)%
∞	6.299(6)%	6.270(3)%	6.368(5)%
Unoptimized	4.58758548%	3.22981197%	4.12127002%

TABLE IV: Critical values for a channel with independent probabilities p of bit flip and phase flips $(p-p^2, p^2, p-p^2)$

Level	[[5,1,3]]	[[7,1,3]]	Special code
0	11.00278644%	11.00278644%	11.00278644%
1	10.94668310%	10.94286393%	11.21042175%
2	10.94728109%	10.95683308%	11.22022045%
3	10.949(1)%	10.960(0)%	11.25155077%
4	10.949(9)%	10.961(5)%	11.247(2)%
5	10.950(4)%	10.962(3)%	11.248(4)%
6	10.950(7)%	10.962(7)%	11.249%
7	10.950(8)%	10.962(9)%	11.249(3)
∞	10.951%	10.963%	11.249(5)%
Unoptimized	7.14780025%	6.45962393%	8.23120201%

which is the same as the channel component $\mathcal{G}^{IIIX}$ for the 4 qubit bit flip code.

The difference between these is

$$\Delta = \mathcal{G}^{IX} - \mathcal{G}^{XI} = [0, 0, YX - XY, ZI - IZ] = [0, 0, x\frac{1-x^2}{2}, x\frac{1-x^2}{2}].$$

We optimize the total 4 qubit code by adding Δ to the unoptimized channel map $[1, 1, x, x]$, and get the optimized map

$$\Omega^{bf_4}[1, 1, x, x] = [1, 1, \frac{3}{2}x - \frac{1}{2}x^3, \frac{3}{2}x - \frac{1}{2}x^3]$$

and then recognizing that $x \to \frac{3}{2}x - \frac{1}{2}x^3$ converges to 1 for $x > 0$, thus iterating this map now yields a useful code for correcting X errors.

Error correcting code capacities In general, the capacity of a code under optimized recovery operators can be estimated by analysis of the entropy of the noise at each level of the code. The situation is found by examining the conditions under which the entropy is equal to some predetermined value, usually 1, at each level of the code.

TABLE V: Critical values for $(p_X, p_Y, p_Z) = (p, 10^{-6}, 10^{-6})$. Note that for the $[[7, 1, 3]]$ code, the entropies are $\frac{1}{2}$.

Level	[[5,1,3]]	[[7,1,3]]	5 bit flip / 5 phase flip
0	49.62410483%	11.00138%	49.62410483%
1	49.64614794%	10.94281%	49.64614908%
2	49.66961046%	10.95678%	49.66961385%

The first two levels of the code were calculated directly, later levels had Monte Carlo calculations. We were able to replicate previous calculations on the average rate of error under concatenation [5]

For 3 different types of noise, the value of p for which the channel Shannon entropy is 1 was calculated for various levels of concatenation. These converge to the channel capacity at the ∞ level. Tab. III is the depolarizing channel. Tab. IV shows results for independent bit and phase flips. Noise dominated by X errors $(p, 10^{-6}, 10^{-6})$ is in Tab. V. In some of the cases, we compare the capacities to the unoptimized thresholds from the channel maps from previous work [4] [6].

The $[[5, 1, 3]]$ code [1] corrects all of the noise that has an entropy of at most 0.993, with the worst case being noise of the form $(p, 0, p)$. It can correct the type of noise in Tab. V in cases where the entropy is greater than 1.

The critical values for the noise $(0, 0, p)$ for the $[[7, 1, 3]]$ code[1] are the same as for the noise $(p - p^2, p^2, p - p^2)$. These correspond to thresholds in the entropy of 0.4988, and 0.9976 respectively, and are the worst and best cases for the threshold in terms of the entropy.

Our best results are obtained with a special code, which is a 5 qubit bit flip code at the first level, a 5 qubit phase flip code at the 2nd and 3rd levels, and the $[[5, 1, 3]]$ code at the higher levels. For the first two levels with the depolarizing noise (Tab. III), the resulting fidelities $f = 1 - 3p$ agree with previous work [2] [7].

This adaptive concatenation greatly outperforms the concatenated channel map method. In Tab. III, it can be seen that for depolarizing noise, the $[[7, 1, 3]]$ code has a capacity of $p = 6.27\%$ instead of $p = 3.2298\%$. For other types of noise and rates of convergence see [3].

We thank K. B. Whaley and M. Sarovar for helpful discussions. We thank the NSF for financial support under ITR Grant No. EIA-0205641.

* Electronic address: jesse@math.berkeley.edu

[1] I. L. Chuang and M. A. Nielsen. *Quantum Computation and Quantum Information*. Cambridge Univ. Press, 2000.

[2] D. P. DiVincenzo, P. W. Shor, and J. Smolin. Quantum channel capacity of very noisy channels. *Phys. Rev. A*, 57(2):830, 1998. `quant-ph/9706061`.

[3] J. Fern. to be published.

[4] J. Fern, J. Kempe, S. Simić, and S. Sastry. Generalized performance of concatenated quantum codes – a dynamical systems approach. *IEEE Trans. Autom. Control*, 51:448, March 2006. `quant-ph/0409084`.

[5] D. Poulin. Optimal and efficient decoding of concatenated quantum block codes. `quant-ph/0606126`, 2006.

[6] B. Rahn, A. C. Doherty, and H. Mabuchi. Exact perfomance of concatenated quantum codes. *Phys. Rev. A*, 66(032304), 2002. `quant-ph/0206061`.

[7] G. Smith and J. Smolin. Degenerate quantum codes for pauli channels. *Phys. Rev. Lett.*, 98(3):030501, 2007. `quant-ph/0604107`.